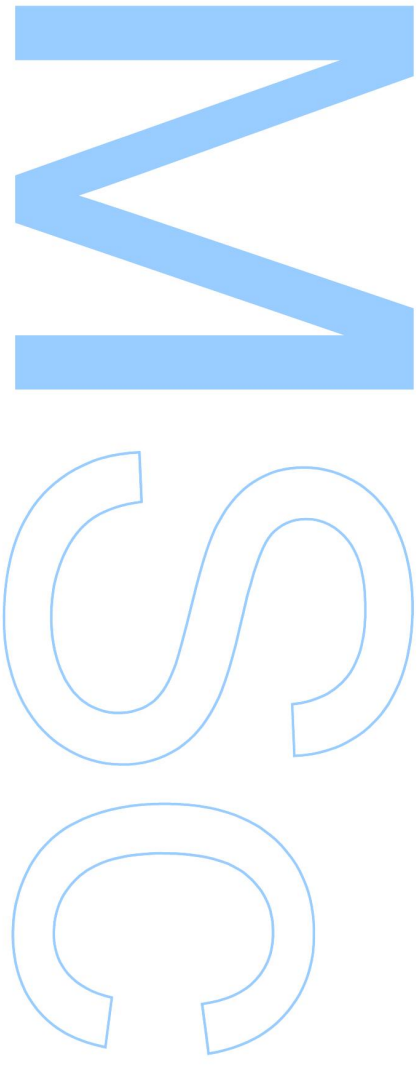
Metrics for Risk Assessment in Data Protection Impact Assessments

Ana Cristina de Melo Carvalho
Mestrado em Segurança Informática
Departamento Ciência de Computadores
2018

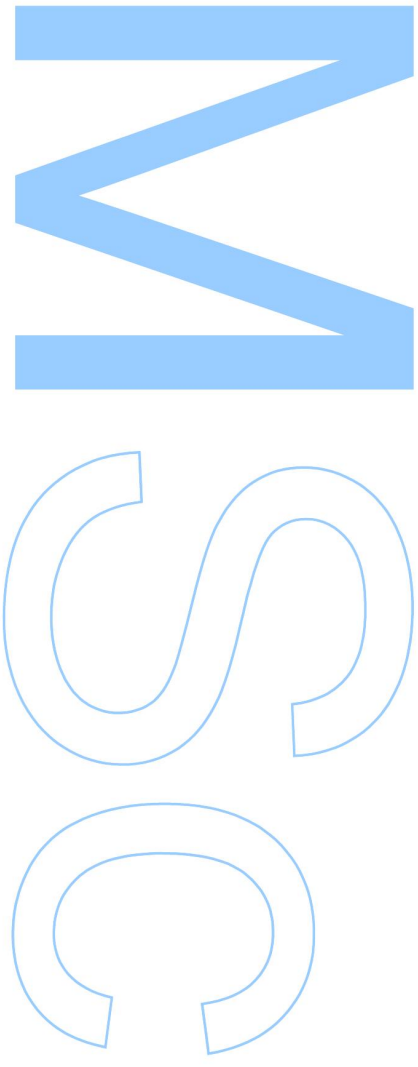
Orientador
Luís Filipe Coelho Antunes, Professor Associado,
Faculdade de Ciências da Universidade do Porto



Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____



Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Professor Luís Antunes for the continuous support of my Master study and research, for all of the opportunities I was given to conduct my research. His guidance, motivation and immense knowledge that always helped me during the research and while writing this thesis.

I would like to thank Professor Rolando Martins for the interest that has always shown for this thesis work and for all the support and ideas given.

I, also, would like to thank Professor Cristina Santos contributions made to this work.

My sincere thanks to all CSIRT.UPORTO team that always supported me greatly and were always willing to help me. I would particularly like to give a special thanks to Eng. Luís Valente for all the support and excellent cooperation and for all of the opportunities that were given by the time I was at U.Porto and during my thesis.

I would also like to thank my colleagues from C3P for their patience and support during the last year, for all suggestions that improved my thesis.

Finally, but not least of all, I would like to thank my family for their wise guidance, support and patience specially during the bad moods. You are always there for me. Finally, there are my friends. We were not only able to support each other in the development of our thesis and work, but also for the well-spent moments of leisure.

Abstract

The new General Data Protection Regulation (GDPR) has delegated more responsibility to the data controllers, who are accountable for the way the data is processed. Thus, in order to minimize the privacy risks, the controllers have to choose the best practices and mechanisms to process data.

Since this risk analysis was not a competence of data controllers, it was the Data Protection Authorities whom developed the risk analysis, the elaboration of this analysis is a challenging task for the controllers, since there are neither established nor professional methodologies with these competencies.

With this work, we intend to facilitate the elaboration of risk analysis and, if necessary, the Data Protection Impact Assessments. In the first instance, describing the existing technological solutions to express online consent in a positive fashion and offering a risk proposal based on the linear combination of the rating of each one of these properties.

Then, by elaborating a semiautomatic assessment that measures the compliance of information systems with the GDPR, from several yes/no questions that are weighted according to the importance of the attribute in relation to the GPDR. In order to elaborate the questions, a mapping of the regulation with ISO 27001 was carried out, and the requirements of some regulatory authorities regarding the security and protection of information systems were studied.

Finally, the agreement among several specialists in the definition of risk was calculated. From the analysis, there is a low agreement between the observers, highlighting that it is not easy to fulfill the requirements of the GDPR and showing that these studies are important when performing a Data Protection Impact Assessment. To overcome the low agreement, we propose the median of the observers' rate.

keywords: GDPR; DPIA; data protection; privacy; consent; information security

Resumo

O novo Regulamento Geral de Proteção de Dados (RGPD) delegou mais responsabilidade aos responsáveis pelo tratamento, sendo estes responsáveis pela forma de como os dados são processados. Assim, para minimizar os riscos de privacidade, os responsáveis devem escolher as melhores práticas e mecanismos para o processamento de dados.

Como essa análise de risco não era uma competência dos responsáveis, eram as Autoridades de Proteção de Dados que efetuavam a análise de risco, a elaboração dessa análise é hoje uma tarefa desafiadora para os responsáveis, uma vez que não existem metodologias estabelecidas nem profissionais com essas competências.

Com este trabalho, pretende-se facilitar a elaboração das análises de risco e, se necessário, as Avaliações de Impacto de Proteção de Dados. Descrevendo-se as soluções tecnológicas existentes para expressar o consentimento online de forma positiva e apresenta-se uma proposta de quantificação do risco.

Em seguida, apresenta-se uma avaliação semiautomática que calcula a conformidade dos sistemas de informação com o RGPD, a partir de diversas questões de resposta sim/não que são valorizadas de acordo com a importância do atributo em relação ao RGPD. Para elaborar as questões foi realizado um mapeamento do regulamento com a ISO 27001, e estudado os requisitos exigidos por algumas autoridades reguladoras quanto à segurança e proteção dos sistemas de informação.

Por fim, calculou-se a concordância entre varios especialista na definição do risco. Da análise, observa-se uma baixa concordância entre os observadores, destacando que não é fácil cumprir os requisitos do RGPD, demostrando que estes estudos são importantes para a realização de uma Avaliação de Impacto de Proteção de Dados. Para superar a baixa concordância, propõe-se a mediana da classificação dos observadores.

keywords: RGPD; AIPD; proteção de dados; privacidade; consentimento; segurança de informação

Contents

List of Tables

List of Figures

Acronyms

AEPD *Agencia Española de Protección de Datos* (Spanish DPA)

APDL *Administração dos Portos do Douro, Leixões e Viana do Castelo, SA*

Art.29 WP Article 29 Working Party

CIA Confidentiality, Integrity and Availability

MP *Município do Porto* (Porto City Hall)

CNIL *Commission Nationale de l'Informatique et des Libertés* (French DPA)

CNPD *Comissão Nacional de Proteção de Dados* (Portuguese DPA)

CRP *Constituição da República Portuguesa* (Constitution of the Portuguese Republic)

DPIA Data Protection Impact Assessment

DPO Data Protection Officer

EDPB European Data Protection Board

EES Economic European Space

eID Electronic Identification

GDPR General Data Protection Regulation

GO Porto Gestão e Obras do Porto, EM

ICO. Information Commissioner's Office

IS Information System

ISMS Information Security Management System

PII Personal Identifiable Information

U.Porto *Universidade do Porto* (University of Porto)

RFID Radio-Frequency Identification

Chapter 1

Introduction

The private life of the individuals, more precisely, the knowledge and study of their personal data is very valuable for States and organizations. Nevertheless, the misapplication of this information might bring limitations to fundamental rights and principles like freedom, equality and personal identity for the individuals.

In the recent years, we witnessed an exponential progression in all technological areas, with special prevalence in computer systems. In parallel to the evolution of online services, technological infrastructures have undergone major changes with the improvement of network connectivity and hardware capabilities. This evolution has been leading society to be more dependent on technology, including social networks, e-commerce, information retrieval, the internet of things among others.

As a result of this technical development and subsequent increasing of data gathering and sharing, there is a need to better protect individuals and their personal data. The actions associated with the usage of applications, including the exchange of information and personal data, can easily be accessible at the network level. Furthermore, during these transactions, all personal data is being reused, linked and analyzed on an unprecedented scale, challenging the privacy of the data subjects.

1.1 Motivation

Unfortunately, most of the personal data processing are not safe. When asked about privacy, most of the individuals associate it with data breaches, social networks exposure or targeted advertising.

It was, consequently, in this environment, that the General Data Protection Regulation (GDPR) appeared, aiming to harmonize data privacy laws across Europe, in order to protect and empower the privacy of all EU citizens and reshape the way organizations approach data protection [10].

Therefore, the GDPR imposes some changes when processing personal data, requiring the implementation of several measures, some of them already present in the previous Directive, ranging from data collection, to its usage, storage, forwarding or sharing and lastly to its destruction, promoting principles of transparency and accountability.

To that end, the organizations must be conscious of the impact that their personal data processing has on the lives of data subjects, identifying and measuring the risks of the data processing, and in some cases extended it to a Data Protection Impact Assessment (DPIA).

Therefore, a DPIA helps to identify privacy issues at an early stage and reduce costs in management time, legal expenses and potential media or public concern by detecting potential risks arising from processing of Personal Identifiable Information (PII). So, a DPIA can be used as an early warning system, informing an organization of the precautions it should take and the tailored safeguards to be taken before further expensive investments [29].

DPIAs, also, helps the organizations to demonstrate the compliance of the assessed object with relevant privacy and data protection requirements [29].

1.2 Proposed Solution

With this work we aim to facilitate the compliance with the GDPR within the organizations, especially regarding the preparation of DPIAs and respective measures to be implemented to reduce the risk.

The term *Risk Analysis*, within the context of information security, can be understood as the "*process that identifies and assesses in a systematic, methodological and repeatable manner the security risks to which the critical business resources of the organizations are subject, enabling the definition of the means by which they can be protected*".

Therefore, risk management allows to determine the precautions to take with regard to the data nature and the risks of the processing, to preserve the data security, thus protecting personal data requires taking "appropriate technical and organizational

measures to ensure a level of security appropriate to the risk" [article 32 of the GDPR] [16, 19].

During the assessment of the privacy impact of processing personal data, all the factors and actors involved in the processing must be counted, from the selection of personal data processors, the forms available to data subjects exercise their rights, the technologies used to process the data, among others.

Such approach allows for objective decision making and the determination of the measures strictly necessary and suitable to the context. It is, however, often, difficult when you are not familiar with those methods, to apply such an approach and to ensure that the required measures have indeed been implemented [16].

Therefore, consent acquisition is one of the factors that influences the DPIA and one of the factors that introduced several alterations to the normal work-flow of institutions. In this thesis we enumerate and assess the existing mechanisms for consent acquisition of data subjects to process their personal data. The purpose of this assessment was to standardize and ensure consistency in DPIAs, and to facilitate the choice of method to obtain consent from data subjects.

Also, in the preparation of the DPIAs it is essential to know the conditions in which the digital data is handled and stored. So, an assessment of the Information Systems (ISs) and the security and privacy measures implemented is also indispensable. Thus, in this thesis, we also developed a semi-automatic DPIA for ISs, that semi-automatically calculates the risk of using the system, according with the input given by the responsible for the system suggesting some corrective measures to minimize the risk. The questions were based on a mapping between GDPR and ISO/IEC 27001, and the security and privacy requirements defined by AEPD and CNIL, two European Data Protection Authorities (DPAs).

1.3 Related Work

During the elaboration of the thesis, some orientations and frameworks related to this topic were studied, some of these orientations were already well established in society as ISO 27001, but others were made available during the last year. In this section we present some of the sources of information that guided us in the development of this work and some similar platforms.

1.3.1 GDPR Guidelines

During the period of adaptation to the GDPR, many entities provided guidance on how organizations could adjust to comply with the regulation. One of the most active entity was the Article 29 Working Party (Article 29 WP) [36], set up by the European Commission at the time of Directive 95/46/EC. The Article 29 WP provided guidelines on the most controversial and abstract issues of the GDPR, explaining in a more concise way what was expected from the organizations to comply with the regulation. Having presented, among others, guidelines on the consent request and how to carry out a privacy impact assessment.

Data Protection Authorities (DPAs), in particular the CNIL, ICO. and AEPD, also presented some guidelines to help controllers and processors to comply with the regulation, especially in the development of DPIAs [7, 12, 25].

Standards already accepted for the protection and security of personally identifiable information (PII) were also used as guide, such as ISO 27001 (Information technology: Security techniques: Information security management systems: Requirements), ISO 29134 (Information technology: Security techniques: Code of practice for personally identifiable information protection), ISO 27018 (Information technology: Security techniques: Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors) and ISO 29100 (Information technology: Security techniques: Privacy framework) [26–29].

1.3.2 GDPR Frameworks

In addition to the guides, some frameworks were also made available, during the last year, for compliance with the GDPR.

As the system available by CNIL to perform DPIAs, in which the controller can characterize each of its treatments and carry out a risk analysis, freely identifying vulnerabilities, threats and corrective measures [13]. Or the test to evaluate the implemented safety measures, made available by the AEPD, which according to closed answers (mostly yes/no) assesses the level of information security and indicates what corrective measures the organization should take [8]. Microsoft, also, made available a semiautomatic risk assessment to assess in which measures the processing complies with the regulation, indicating the corrective measures to be taken in order to reduce the risk [34].

Other DPAs also provided small closed-ended questionnaires so that organizations could assess their level of compliance with the GDPR, and templates of official documents, such as DPIAs, contracts, clauses, requests for consent, among others.

1.4 Contributions: Real World and Scientific

The proposal of DPIA for ISs was employed in an online platform (TekPrivacy) that helps organizations, specially Data Protection Officers (DPO) to guarantee the compliance with GDPR. And it has been, already, implemented in some national organizations like GO Porto, APDL and CMP, in the process of compliance to the GDPR.

During this project development, a short paper was accepted by “Privacy Security Trust 2018” to be presented and published at the conference "16th Annual Conference on Privacy, Security and Trust", and indexed with IEEE Xplore Digital Library [11].

1.5 Outline of the Thesis

This work is divided into 6 Chapters. The current chapter presents an introduction of the work of this dissertation.

The next chapters of this document address the following topics:

In Chapter 2, the legal framework for protection of personal data in National reality is discussed, as well as, the role of Article 29 Working Party and the GDPR requirements that imply a "revisiting" of technology. One of the requirements that, so far, has not yet been technically duly defined is the way to request for an auditable and explicit consent, which is discussed in Chapter 3. In Chapter 4, a proposal of DPIAs for ISs is presented, where, an ISO mapping with the RGPD was made. Therefore, Chapter 5 is where we present the assessment of the proposal of DPIA for ISs at CMP, APLD and GO Porto, and explain the environment created to that end. Finally, the Chapter 6 presents some final remarks and lays the ground for future work.

Chapter 2

Legislation

The protection of personal data has an important role in the defense of individuals rights, since it does not only protect privacy, but also helps the protection of principles like freedom, equality and personal identity.

So, the right to respect for private and family life has been established since 1950, in the Universal Declaration of Human Rights in article 12. "*No one shall be subjected to arbitrary interference with his privacy, family, domicile or correspondence, or attacks on honor and reputation, and against such interference or attacks, everyone has the right of the protection of the law.*" [2].

This right is also determined in the International Covenant on Civil and Political Rights in Article 17 "*No. 1 - No one shall be subjected to arbitrary or unlawful interference with his private life, family, home or correspondence, or unlawful attacks on honor and reputation; of the law against such interventions or such attacks.*" [4], and the European Convention on Human Rights in article 8 "*No. 1 - Any person has the right to respect for his or her private and family life, home address and correspondence.*" [3].

This first big emphasis on the principle of reservation of private life was due to World War II, which, due to the great impact of the invasion of privacy committed by the Nazis and the consequent result, it became indispensable for States to limit the access and usage of personal information of individuals.

Furthermore, each State dictates rights protecting the personal data and privacy of citizens, according to each national reality.

2.1 GDPR in the National Reality

2.1.1 Protection of Private Life

In the Portuguese law, the right to privacy of private life, which is embodied in the rights of personality, was established in the Constitution of the Portuguese Republic of 1976 (CRP) (in the initial version) in the article 26 "*1. Everyone is accorded the rights to personal identity, to the development of personality, to civil capacity, to citizenship, to a good name and reputation, to their image, to speak out, to protect the privacy of their personal and family life, and to legal protection against any form of discrimination. 2. The law shall lay down effective guarantees against the improper procurement and misuse of information concerning persons and families and its procurement or use contrary to human dignity. 3. The law shall guarantee the personal dignity and genetic identity of the human person, particularly in the creation, development and use of technologies and in scientific experimentation.*" [6].

The processing of personal data might induce more information about the private life of an individual and that information can correlate data that, as a rule, would not be related. Thus, as we lose control of the result of the processing, the content generated may be liable to harm the data subjects. So, it is of utmost importance that the data processing is carried out in a conscious way and by people with the capacity to assess the impact of its use on individuals' lives. As we are addressing sensitive personal information, the ownership of the information is guaranteed to the own individual, so that the individuals can protect their information from access, treatment and sharing, and respective consequences of their processing. The possessiveness can be observed by the fact that, excepting special cases, only with the consent of the owner of the information (data subject), personal data processing can be performed.

2.1.2 Data Protection

As a result of the development of technologies and the sharing of information on a regular basis, such as e-commerce, internet of things and big data, there is an increasing need to better protect individuals privacy and their personal data.

Thus, probably predicting the technological developments and to control this expanding industry, the Constitution of the Portuguese Republic (CRP) covers the protection of personal data in the article 35 (Use of information technology) "*1. Every citizen has*

the right of access to all computerized data that concern him, which he may require to be corrected and updated, and the right to be informed of the purpose for which they are intended, as laid down by law. 2. The law shall define the concept of personal data, together with the terms and conditions applicable to its automatized treatment and its linkage, transmission and use, and shall guarantee its protection, particularly by means of an independent administrative entity. 3. Information technology may not be used to treat data concerning philosophical or political convictions, party or trade union affiliations, religious faith, private life or ethnic origins, save with the express consent of the data subject, or with an authorization provided for by law and with guarantees of non- discrimination, or for the purpose of processing statistical data that are not individually identifiable. 4. Third-party access to personal data is prohibited, save in exceptional cases provided for by law. 6. Everyone is guaranteed free access to public-use information technology networks. The law shall define the regime governing cross-border data flows, and the appropriate means for protecting both personal data and other data whose safeguarding is justified in the national interest. 7. Personal data contained in manual files enjoy the same protection as that provided for in the previous paragraphs, as laid down by law." [6].

Alongside the principals defined in the CRP, the Personal Data Protection Act (Law 67/98, of 26 October) [5], based on the Directive 95/46/EC of the European Parliament and the Council, dated October 24, 1995 [18], was implemented in 1998 and it was the law in force until 25th of May of 2018. This law aimed to guarantee the protection of the fundamental freedoms and rights of natural individuals, in particular privacy, in relation to the processing of personal data. The Personal Data Protection Act defined 'personal data' as "*... any information of any type, irrespective of the type of medium involved, including sound and image, relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an indication number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;*" [5] and the 'processing' of personal data as "*...any operation or set of operations which is performed upon personal data, whether wholly or partly by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;*" [5].

However, with the technological evolution, where more and more IT applications are being used, data being collected, and the States themselves implemented online services, ensuring the cybersecurity and protection of the personal data is crucial. What was no

longer sufficiently guaranteed by the Data Protection Directive 95/46/EC, not only because of the obvious technological progress and the need to process data, but also because it is a directive and consequently not all EU countries met all the factors defined by it, since each member State adapted the Directive according to National reality.

2.1.3 General Data Protection Regulation

Therefore, the General Regulation on Data Protection (Regulation (EU) of the European Parliament and of the Council 2016/679 of 27 April 2016) was designed to harmonize data privacy laws across Europe to protect all private data of EU citizens, and to reshape how the organizations of the Union address data privacy [10] laying down rules on the protection of individuals with regard to the processing of personal data and the free circulation of such data [Art. 1, paragraph 1 of the GDPR] [19].

However, some changes imposed by the GDPR in the approach of processing data enforced the implementation of several measures. Whereas, some of these concepts are already uniformed and applied in some organizations, others are a new concept that the institutions still have some questions on how to implement them.

Some of the changes present in the GDPR encompasses the increase of the territorial scope, the improvement of the definition of consent, the definition of special data, increase of the rights of data subjects, control of automated decisions, mandatory implementation of the concept Privacy by Design, the definition of Data Protection Officer and the values of the fines [10].

In order to help the compliance with GDPR, the Article 29 Working Party [36] has elaborated articles, about some issues addressed in the regulation, exemplifying and explaining how to be *GDPR compliant*.

Even so, the measures imposed have such a large impact on organizations that this regulation had a two-year waiting period, between the entry into force and beginning to be applied, to enable organizations to adapt to it.

2.2 Consent According with GDPR

One of the changes imposed by the GDPR is related to the way consent is expressed from the data subject, serving as a legal basis for the the processing of his personal data. This is one of the main focus of this thesis.

Consent plays an important moral role, with the capacity of transforming the normative expectations that are hold between people and groups, whether directly or through various institutional arrangements, if properly given. Therefore, consent can be compared with a proprietary gate that one opens to allow another's access, which would be impermissible absent the act of voluntary opening such gate [35].

2.2.1 Legal Aspects

Given its importance for the technical aspects considerations, in this subsection, we summarize the list and discussion of legal requirements presented in the guidelines of the Art. 29 Working Party on Consent.

In the GDPR, Articles are binding and Recitals are used for interpretative purposes, that is why, in this work, we try to use the recitals. By doing this, some sub-components of the elements of the definition of personal data (informed, specific, freely given and unambiguous) are *highlighted*. For example, explicit consent is only needed in certain cases, such as for special categories of data.

Notice that, in some cases Recitals are more fundamental than the binding Articles. In particular, Recital 43 stating that for the freely given requirement the provision of the service should not be made conditional upon consent. However, in the binding parts so Article 7(4) the words *utmost account* are used.

According to the regulation, *"consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. This could include ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity should not therefore constitute consent. Consent should cover all processing activities carried out for the same purpose*

or purposes. When the processing has multiple purposes, consent should be given for all of them. If the data subject's consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided."[Recital 32] [19].

Beyond the aforementioned description, the GDPR also defines that *"where processing is based on the data subject's consent, the controller should be able to demonstrate that the data subject has given consent to the processing operation."*[Recital 42] [19].

Therefore, in order to comply with the regulation, consent expression must have the follow characteristics:

Freely given [Recitals 32, 42, 43 & Article 7 nº 4]
Data subject must have a real choice and control. If the data subject feels obliged to consent or that the non-consent leads to negative consequences, i.e., if the consent is described as a non-negotiable requirement, then the consent is not considered valid [38].

Granular [Recital 43]
Granular is a specification of *Freely Given* but considering the relevance and the novelty of this characteristic we decided to highlight it. Therefore, consent is not considered freely given if the data subject thinks/assumes that all purposes are *all in one*, wherefore data subjects should be free to choose which purposes they accept, rather than having to consent to all processing purposes [38].

Specific [Recital 32 & Article 7 nº2]
This characteristic is tightly coupled with the previous one and aims to ensure transparency and control from data subject [38]. Data subject must be informed of all the purposes related to the data processing, and he/she must choose in relation to each one of them, and to that end. As such the information given must be clearly segregated [38] regarding each purpose.

Informed [Recitals 32, 42 & Article 7 nº2]
Data subject must be informed about the processing operations before the consent is presented. The information must be given in a clear and plain language, where statements filled with legal jargon cannot be used and information relevant to make informed decisions cannot be hidden [38]. If the target audience includes data subjects

that are underage, the provided information must be understandable by them [38].

Unambiguous indication of wishes [Recital 32]
Data subject must give a statement or a clear affirmative act that consents the data processing to the different purposes, which means that the consent always must be given through an active action or declaration from the data subjects part [38].

In some cases, all these characteristics are still not enough, and explicit consent by the data subject is required. Some examples are: processing of special categories of data; communication of personal data to an organization or country that does not guarantee the adequate safeguards; or when automated individual decision-making is realized [19, 38].

Explicit [Recitals 51, 71 and Article 9 nº 2 al. a)]
For the consent be explicit, the data subject must give a positive statement of consent, to expressly consent in a written statement the processing operations or transfer of the personal data, since an oral statement might not be audit-capable [38].

In addition to the characteristics already indicated, there are still two additional complements that the consent must contain to be considered valid, namely:

Auditable (demonstrate consent) [Recital 42]
Controllers must be able to demonstrate that the data subject in a given case has consented and that the consent was given in a valid way, linking the consent to the processing operations [38].

Withdrawal [Articles 7 nº 3]
Data subject must be informed on how to withdrawal the consent before the consent is given, as well as any other subject rights. The withdrawal of consent must be as easy as the act of giving it [38].

2.2.2 Technical Aspects

The expression of positive consent is clearly one of the new aspects with the greatest impact on the implementation of the GDPR, due to some technical requirements that have not yet been solved. Since the consent criteria are technically demanding, particular attention should be paid to the form of application, so as not to compromise the usability of the application. As the final user may quit from using the application due to its low usability on the request for a complex consent.

Below, it is presented a technical orientation on the different characteristics identified in the previous section and on how they should be applied.

Freely given

To express consent, the data subject should know of all the implications of the processing operations and cannot be under any kind of coercion, which is impossible to prove, if the person is not visible (typical in an *online* environment).

Data subject cannot feel that by refusing to consent a second purpose, the application cannot be used. Unless processing is required to provide the service, consent is only freely given if the use of an application is not conditional upon consent.

Granular

When the legal base to process personal data is the consent, a previous request of consent to data's access and processing operations should be asked. Moreover, different purposes, e.g., channel used to communicate, processing operations objective and data to be collected, must be identified and segregated in order to allow the data subject to only select the ones he wants.

Specific

The information related to the data processing operations presented in the consent request must be segregated from the rest of the information, the consent requests must be granular and the purpose must be specified [38]. Thus, data subjects must be informed about the intended purposes for the data processing, as such there must be opt-in options for each purpose, with the appropriate information [38].

Informed

The consent must describe the controller's identity with the purposes for each processing operation, the data to be collected, the existence of the subject's rights and how to apply them, activities performed in the processing operations and whether there is data transfer to third countries in the absence of an adequacy decision and appropriate safeguards. All this information should be given before obtaining the consent from data subject.

Unambiguous indication of wishes

Data subject must give a statement or a clear affirmative act that consents the data processing operations. That way, pre-selected options or out-out requests are not accepted, and the data subject must be the one who expresses the consent through a written or (a recorded) oral statement that selects the options being consented, among others [38].

Explicit

The consent must show that it was given by the data subject and clearly identify what was consented and that it had a real action from the data subject to consent the purpose. The explicit consent must also be able to be verified in order to guarantee its validity and that the data subject authorized the use of his/her data for each purpose.

Auditable

The records and purposes of the consent requests must be stored and they must prove that data subject gave the consent. It also has feature the date of submission of the consent, the way it was given, the work-flow that originated it, the information that was given at the time of the request and finally, what data the subject consented to.

Withdrawal

The consent mechanism must indicate also how to withdraw consent at any given time. The withdraw procedure must be as easy as it was to give consent in the first place. Subsequently, the means to withdraw the consent should be presented in the same way as initially the consent was given [38], in particular consent should be revocable for selected purposes.

In fact, the industry has already implemented some methods to minimize the usability

issues regarding the collection of the consent, however, it is still categorized as an obstacle to overall service usability [15]. This problem is further exacerbated due to the fact that it must be auditable and, in some cases, extended to explicit.

In addition, it is also advisable to refresh the consent at appropriate intervals, in order to guarantee the compliance with the wishes of data subjects and remind them of associated data processing operations.

Furthermore, considering the sensibility of children personal data and the required safeguards concerned with their privacy rights, a set of specific conditions must be met in the collection of their consent. More specifically, in order to collect consent for purposes of marketing, creating personality, definition of profiles and collection of personal data with regard to children when using services offered directly to a child, the consent from the holder of parental responsibility must be provided. However, it is not necessary to do so in the context of preventive or counselling services offered directly to a child. Although, the information provided to children should be in presented in a clear and plain language so it can be easily understood by them [19].

In these cases, the consent support mechanism must be specialized so that the request for consent is forwarded to the holder of parental responsibility, so that consent is given to the processing of the child's personal data. Although the regulation argues that this definition of a child is for children under 16 years. This allows each State Member to adapt to its context, thus allowing this age to decrease to 13 years.

2.3 Article 29 Working Party

The Working Party responsible for the Protection of Individuals with regard to the Processing of Personal Data, commonly known as the Article 29 Working Party (Art.29 WP), is an independent EU advisory body on data protection and privacy, composed by representatives from all EU Member States, the European Data Protection Supervisor and the representative of the European Commission [1].

Art.29 WP was conventionalized by Article 29 of Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the dissemination of such data. Its functions are defined in Article 30 of that Directive and in Article 15 of Directive 2002/58/EC on privacy and electronic communications [1].

Art.29 WP examined any question concerning the application of the measures adopted

under the Directive in order to contribute to the standardization of the application of those measures, giving an opinion on the level of protection in the Community and in third countries to the Commission, advising the Commission on any additional or specific measures to safeguard the rights and freedoms of the individuals with regard to the processing of personal data and any other proposals for Community measures affecting such rights and freedoms and issues an opinion on codes conduct at Community level [1].

The Group, on its own initiative, makes recommendations on all matters relating to privacy and data protection in the European Community. Art.29 WP conducted online consultations on data protection issues related to intellectual property rights, RFID, video surveillance, binding corporate rules, electronic health records and, more recently, the protection of children's personal data. It has issued a number of opinions on subjects such as the transfer of information from travelers' personal identification records to the US, the introduction of biometric data in passports and visas, the transfer of financial information to the US, the introduction of data retention requirements across the EU, a draft of the Commission decision on standard contractual clauses for the transfer of personal data to processors established in third countries pursuant to the Directive 95/46/EC and on proposals to amend the Privacy Directive and Electronic Communications (Directive 2002/58/EC) [1].

The opinions and recommendations of Art.29 WP are sent to the European Commission and to the Committee set up to assist Article 31 of the European Commission Directive 95/46/EC that informs Art.29 WP of the measures it has taken in response to its opinions and recommendations through a public report. This report is also forwarded to the European Parliament and the Council. Art.29 WP is as well responsible for publishing an annual report highlighting the evolution of data protection across the EU and in each EU country and the European Economic Area. This annual report is also transmitted to the Commission, the European Parliament and the Council [1].

However, with the application of GDPR the Art.29 WP will be replaced by the European Data Protection Board (EDPB) [19]. Nevertheless, until its replacement the Art.29 WP is providing expert advice to the States regarding data protection and promoting the consistent application of GDPR, like was given under the Directive 95/46/EC.

2.4 GDPR Requirements that Imply a "Revisiting" of Technology

As technology has advanced, as well as the physical to digital process transfer, the mechanisms should also have advanced to ensure the same levels of digital security and privacy found in the physical.

Although most of the measures are procedural, the GDPR also imposes the organizations to take extra care on the conditions the data is handled technically, to re-calibrate the security and privacy level between physical and digital environments, requiring the implementation and alteration of the process of diverse technological tools, implying the "revising" of current technological mechanisms.

Most of these changes are related to the concepts introduced by privacy by design and privacy by default written into Article 25, which although not new, they have now become mandatory, implying that any action by an organization involving the processing of personal data must be made with data protection and privacy in mind at all stages (Privacy by Design), and once a product or service has been released to the public, stiffer privacy settings should be applied by default, without any manual end user input (Privacy by Default).

One of the key aspects can be considered the procedures to collect consent, as they shall be audit-capable and enable data subjects to express their consent according with the GDPR requirements; the methods to execute the right to be forgotten, that should guarantee that when the data is removed, its reconstruction should be impossible; the auditability of ISs; and the anonymity and pseudonymity of the data when the identification of the individuals is not requested.

These changes in technologies are implemented by organizations to ensure that they are in compliance with the fundamental principles and requirements of GDPR and those alterations are part of the focus on accountability.

Chapter 3

Expressing Explicit and Auditable Consent

There are several legal grounds to allow the possessing of personal data, being one of them the consent of the data subject. Consent gives data subjects control over his personal data, processed or not [38]. Thus, if consent is used correctly, it is a powerful tool that gives data subjects control over the processing of their personal data, otherwise, which is the prevalent case, control of the data subject becomes illusory and consent constitutes an inadequate basis for data processing [37].

Given the power that consent has on the processing of personal data, the way to request consent of the data subject on the processing of his/her personal data was revised in GDPR, becoming one of the key changes of the regulation.

Consent Historical Context

Consent plays an important moral role, with the capacity of transforming the normative expectations that are hold between people and groups, whether directly or through various institutional arrangements, if appropriately given. Therefore, consent can be compared with a proprietary gate that one opens to allow another's access, which would be impermissible absent the act of voluntary opening such gate [35].

The history of the informed consent is rooted in multiple disciplines and social context, including those of the health professionals, law, the social and behavioral sciences, and moral philosophy [20]. Being interpreted as a form of respect for autonomy and individuality of the individuals. Consequently, informed consent is interpreted as a

moral principle of respect for autonomy [20].

With the technological evolution, a further scope of application of informed consent was found, arising as an issue of paramount importance within the scope of the protection of privacy and of the processing of personal data [32]. This application is even more challenging, since it is found in a different (digital) environment, and thus, it is necessary to find new ways to adjust the consent of individuals.

3.1 Existing Technologies to Express Consent

In the present-day, a lot of solutions are being implemented for organizations to request data subjects to express their consent to the processing of their personal data. Although, not all of them fully meet the requirements imposed by GDPR.

The majority of consent mechanisms applied until now, fail to meet the requirements of GDPR, because they were given in an implicit or opt-out way. Notice that, the notion of implicit consent was not defined and foreseen by the Directive, neither the opt-out practice. The implicit mode defines that the data subject feels that the consent must be given in order to use the product - most of the times appearing with only one button to agree -. The opt-out mean implies that the data subject must remove the consent, being pre-assumed that the consent is given - most of the times the way to withdraw the consent is omitted -.

With the increasing awareness of the measures imposed by GDPR, new proposals of mechanisms to give consent started to appear, mainly in a way where the description of the processing operations is clearly presented to the individual and identifies for what and whom the data will be accessed. So, the data subject has the option to agree or disagree with each processing operations and respective data to be collected, which means that those consent requests comply with the request of a regular consent.

Nevertheless, not all of them can be classified as being explicit. For a consent be explicit, the records related to the consent must prove that it was the data subject that gave consent, that existed a positive action to consent, and this must be unequivocal, expressing exactly the wishes of the data subject. These characteristics must also be followed with the audit-capable property, that features the date of submission of the consent, how it was given and the information that was given in the request.

3.1.1 Existing Technologies

Advanced Electronic Signature

The consent process includes the form with a field to add the digital signature and a submission button. This process can be done online and does not need any equipment since the signature can be made from a certificate that is stored on the computer, minimizing usability issues identified in the *Signature - Present*, as we will see below.

The digital signature with a certificate that is not validated by a competent identity, a signature with an advanced electronic certificate, it must be uniquely capable of identifying and linking its signatory and guarantees the integrity of the document. So, a qualified electronic signature is an advanced electronic signature with a digital certificate that has been encrypted by a secure signature creation device.

Code Sent by email

This consent method includes the form with a submission button and two fields, one to add the email, where a confirmation code will be sent, and other to place the code received. After getting the consent form completed, the data subject submits the code received by email, and sends the form.

Code Sent by SMS

Similar to the previous one, the consent system includes the form with a submission button and two fields, one to add the phone number where a confirmation code will be sent, and other to place the code. After getting the consent form completed, the data subject submits the code received by SMS, and sends the form.

Code to Access the Consent Given

Similar to two-factor authentication by SMS or email, but in this case the code is not used to submit the consent, instead it is to confirm the consent given.

Thus, the consent form has a submission button and a field to add a phone number or an email to where the code will be sent after the data subject submits the consent. Therefore, the data subject can access the form to verify if it is in accordance with the expected and change whenever appropriate.

Confirmation Button of the Consent by Email

The consent form, in this mechanism, has a submission button and a field to add an email address. So, after the form is submitted, the data subject receives an email with all the parameters consented, so that the data subject can confirm by pressing the button that leads to a link.

Consent Given Sent by Email

The consent form, in this mechanism, has a submission button and a field to add an email address. Therefore, after the form is submitted, the data subject receives an email with all the parameters consented, so that the data subject answers it to confirm.

Digital Mobile Key with SMS

The Digital Mobile Key is a mechanism of authentication that allows the association of a mobile phone number with the civil identification number of a citizen, this method is provided by some countries, such as in Portugal.

Thus, the consent system includes the form with a submission button and three fields, one to add the phone number where a confirmation code will be sent, other to add the PIN of Digital Mobile Key and other to place the code. After getting the consent form completed, the data subject submits the code received by SMS, and sends the form.

Digital Mobile Key with Smartphone Application

The application Digital Mobile Key is an alternative to the mechanism previously mentioned, being this time, the code sent by a push notification to the smartphone that is related to the civil identification number. With this mechanism is also possible to generate new codes and control the live time of each one. This method is similar with Google Authenticator.

Thus, the consent system includes the form with a submission button and three fields, one to add the phone number, other to add the PIN of Digital Mobile Key and other to place the code generated in the application. After getting the consent form completed, the data subject submits the code, and sends the form.

Email from Data Subject

The data subject sends by email to the controller of the processing operations with the form provided by the controller, indicating what the data subject consents.

Login Authentication

This mechanism implies that the data subject is already authenticated when the consent is completed and submitted. Which by default is a problem, because in order to give consent, the individual must already have an account and consequently already must be using the application.

The consent form has a submission button, the record of the consent is saved on the user account after this is sent, thus it is documented.

Qualified Electronic Signature

Digital signatures are like electronic "fingerprints." In the form of a coded message, the digital signature securely associates a subscriber with a document in a recorded transaction, using a standard and accepted format it provides the highest level on both security and universal acceptance [17].

Therefore, the certificate present in the qualified electronic signature, like national electronic identification (eID) schemes, is issued by a qualified trust service provider, that attests the authenticity of the electronic signature to serve as proof of the identity of the signatory.

Accordingly, the consent process includes the form with a field to add the digital signature and a submission button. This process can be made online, which minimizes the usability issues identified in the previous mechanism.

The citizen ID card managed by some governments is an example of a qualified electronic signature. However, in most cases obligates the possession of the necessary equipment and software to use it, limiting the data subjects' universe.

Signature - Present

A signature is used to permanently appended a single person, in an indelible manner, to a document. It then can be used as physical evidence of that person's personal testimony and certification of the signed content.

Thus, at the time of giving consent, it is demanded the physical presence of data subjects to sign and date the form. After the consent is given from data subject, a copy of the document is given to data subject and signed by both parties.

Submission of a Document Proving Identity
This consent system includes a form with submission button and a field to submit a document that proves the identity of the data subject. The data subject completes the form, submits a document and sends the form.

Voice Call
This consent system is frequently used by the marketing companies, where companies come in contact with the data subject through a voice call. In this process, a call is recorded and the consent of the data subjects for different purposes is recorded. During the call, all the necessary information for an informed consent is given and any questions presented by the data subject can be resolved at the time.

3.1.2 Compliance of the Technologies with GDPR

The necessary mappings between the mechanisms identified in this section and the previously defined characteristics that the consent must meet are shown in Table 3.1. As a recap, the characteristics include Freely given (F), Granular (G), Specific (S), Informed (I), Unambiguous indication of wishes (Un), Auditable (A), Withdrawal (W) and Explicit (E). Furthermore, it also shows if the mechanism guarantees the Data Minimization (DM) principle, and whenever the information can be done Electronically (El) or it must be in person.

In order to carry out the analysis of compliance with GDPR, it is assumed that, for each mechanism, the maximum possible completeness with the identified characteristics of the consent are implemented, so we assess the best possible scenario for each mechanism by itself. Thus, it is analyzed at its best, how the system complies with the GDPR.

Thus, when analyzing Table 3.1 we can verify that all mechanisms can only guarantee the completeness of the characteristics granularity, specificity, information, unambiguous indication of wishes and withdraw of the consent as defined by GDPR. The remaining can only be accomplished by some methods, either totally or partially, and in some cases, it is still needed the combination of other actions. Those cases are explained below.

Freely Given
The mechanisms *Voice Call* and *Signature - Present* can contemplate all the require-

Table 3.1: Features accomplished by the consent mechanisms identified.

Mechanism	F	G	S	I	Un	A	W	E	DM	El
Advanced Electronic Signature	✘	✓	✓	✓	✓	✘	✓	✓	✘	✓
Code Sent by Email	✘	✓	✓	✓	✓	✘	✓	✘	✓	✓
Code Sent by SMS	✘	✓	✓	✓	✓	✘	✓	✓	✘	✓
Code to Access the Consent Given	✘	✓	✓	✓	✓	✘	✓	✘	✓	✓
Confirmation Button by Email	✘	✓	✓	✓	✓	✘	✓	✘	✓	✓
Consent Given Sent by Email	✘	✓	✓	✓	✓	✘	✓	✓	✓	✓
Digital Mobile Key with Application	✘	✓	✓	✓	✓	✘	✓	✘	✘	✓
Digital Mobile Key with SMS	✘	✓	✓	✓	✓	✘	✓	✓	✘	✓
Email from the data subject	✘	✓	✓	✓	✓	✘	✓	✓	✓	✓
Login Authentication	✘	✓	✓	✓	✓	✘	✓	✗	✓	✓
Qualified Electronic Signature	✘	✓	✓	✓	✓	✘	✓	✓	✘	✓
Signature - Present	✓	✓	✓	✓	✓	✓	✓	✓*	✘	✗
Submission of a Document	✘	✓	✓	✓	✓	✘	✓	✘	✗	✓
Voice Call	✓	✓	✓	✓	✓	✓	✓	✓	✗	✘

F = **F**reely given;
G = **G**ranular;
S = **S**pecific;
I = **I**nformed;
Un = **Un**ambiguous indication of wishes;
A = **A**uditable;
W = **W**ithdraw;
E = **E**xplicit;
DM = **D**ata **M**inimization;
El = **El**ectronically.

✓= Accomplished;
✘= Partially accomplished;
✗= Not accomplished.
* = Needs an extra action.

ments in order to the consent being freely given. However, in the presence of other people during the action of consent can be considered coercive.

Although, in the remaining mechanisms we cannot observe if the data was given under coercion, so we consider that they only partially comply with the requirements, since they are able to give an explanation of all the implications of the processing operations.

Auditable

Likewise, the mechanisms *Signature - Present* and *Voice Call* are the only ones that entirely comply with the conditions defined for being considered auditable as defined by GDPR.

On the other hand, although the method *Qualified Electronic Signature* guarantees the identity of data subject, it cannot guarantee that the data subject is signing the correct form, but only that what was signed is irrevocable. This problem also occurs with an *Advanced Electronic Signature*, but in the latter, the identity of the data subject is not guaranteed.

Finally, the rest of the mechanisms still do not have a way to ensure that the document that is being accepted is the intended one. It is also necessary that all objects used to attain consent must be signed shortly after being used in order to date the consent.

Every digital consent's models should permit the consent to be printed after given, in order to provide proof to the data subject. However, this option would cause a storage problem.

Explicit

The electronic signatures guarantee the non-repudiation of the signed documents/forms, so they are the only methods that comply with the legal requirements just by their specifications.

The *Signature - Present* can only be explicit if it is guaranteed that the document where consent has been given, no other new authorizations can be added. For that, the data subject must discard the options that were not consented, like drawing a plus sign in the not wanted boxes, making it impossible to add consent in those options.

The rest of the methods might demonstrate the intention of data subject to give consent, but they can suffer spoofing and man-in-the-middle attacks, which means that the identity of the data subject cannot be guaranteed.

However, although the mechanisms *Code sent by Email* and *Code sent by SMS* might suffer from spoofing attacks, it is possible to validate the items consented, by generating a code that have a predefined part that identifies which items were selected. These methods allow greater validation on the part of the user since it can validate the first part of the code, confirming if its choices coincide.

It should be mentioned that the method *Submission of a Document*, which might seem

relatively secure, it is in fact not safe, as the document can be easily forged, so in turn, it can also suffer from spoofing attacks.

Data Minimization

The consent mechanisms themselves should not be intrusive and create a higher risk than the processing operation leading to consent. Thus, although those mechanisms cannot guarantee the anonymity of the data subject, most of them are not very intrusive, most of them only requesting an email address, being the data subjects able to hide their real identity. The trad-off between auditability and data minimization should be balanced by the data sensibility.

The *Login Authentication*, according to the registration form of the platform, may be the least intrusive, allowing data subject to create an account anonymously.

The mechanisms that collect phone numbers are more intrusive than the previous mentioned, since in most cases the data subject uses the personal number.

A signature, digital or paper collects more data, but it is the most reliable method to guarantee the identity of the data subject, so by the proportionality criteria this method is a good choice if it is intended to obtain the real identity of the data subject.

Electronic certificates can have pseudonyms, which provides a method so that at the level of authentication, the data subject identity cannot be crossed by different service providers, and allows the signature from pseudonyms, this feature is already in use on German identification cards [9].

Nevertheless, there are some mechanisms more intrusive like the *Voice Call* that collects a biometric standard, voice, so this mechanism can be too intrusive in order to collect a consent. Depending on the information contained in the document, the *Submission of a Document* may also be overly intrusive for the intended purpose, taking into account that it is also not a very reliable mechanism.

3.2 Risk Analysis

The term *Risk Analysis*, inserted in the context of personal data protection, can be understood as the "process that identifies and evaluates in a systematic, methodological and repeatable manner the risks on personal data to which the critical business resources of the organizations are subject, enabling the definition of the means by which they

can be protected".

With regard to the processing operations under GDPR, a Data Protection Impact Assessment (DPIA) is advised, and in some cases required, in order to assist the organizations in identifying and minimizing the privacy risks of new projects or policies [24].

Therefore, for the elaboration of the DPIA, the data controller must prepare a risk analyses of the processing operations. Being, one of the parameters that must be assessed, the way of asking for consent, thus this work was developed in order to help in the elaboration of DPIAs, particularly this chapter elaborates an assessment of the consent methodologies.

To carry out the risk analysis, seven raters were asked to evaluate according to their experience the different parameters in the various mechanisms. For the selection of raters, their experience and diversity in the scope of activities were taken into account, and then selected according to convenience. Thus, the raters experience are: one MSc student, one PhD student, one Assistant Professor with more than 10 years' experience, one Associate Professor with more than 20 years' experience, one security expert with more than 10 years' experience, and finally two technical privacy professionals working in two different data protection authorities with more than 10 years' experience.

The mechanisms for consent were assessed in their compliance regarding the requirements defined in GDPR and their usability and threats, in a rating scale of 1(very bad), 2(bad), 3(good) and finally 4(very good) for the purpose of elaborating their Risk Analysis, as described in the table 3.2. Because the characteristics granular, specific, informed, unambiguous indication of wishes and withdraw can always be achieved by all methods, they were not considered in this analysis.

In order to calculate the risk, weights were used to differentiate the most important characteristics, with Trustworthy (T) - which quantifies the confidence of the mechanism in accomplishing the expected, the probability of attacks that the mechanism can undergo and the level of ease in carrying out the attack - being the most important - since is the one that levels the confidence in the model -, followed by the Explicitness (E) and Auditability (A) - because they are those that have a more complex scope -, then Freely given (F) and finally Usability (U) - since is out of the scope of the regulation -. Thus, the formula used was

$$Risk(\%) = \frac{MAX\_RISK - (F + 1,5A + 1,5E + 0,5U + 2T)}{MAX\_RISK} \times 100.$$

Notice that the risk calculation expression was created based on the authors experience,

so it should not be taken as a standard before further evaluation and validation.

Table 3.2: Risk analysis of the consent mechanisms identified.

#	Mechanism	F	A	E	U	T	Risk
1	**Qualified Electronic Signature**	4	4	4	2	4	12%
2	**Signature - Present**	4	4	4	1	4	13%
3	**Digital Mobile Key with Application**	3	3	3	3	3	23%
4	**Digital Mobile Key with SMS**	3	3	4	3	3	23%
5	**Code Sent by SMS**	3	2	4	4	3	27%
6	**Advanced Electronic Signature**	2	3	4	2	3	31%
7	**Code to Access the Consent Given**	3	2	3	3	2	31%
8	**Voice Call**	3	3	3	3	2	33%
9	**Email from the Data Subject**	2	3	4	3	2	35%
10	**Code Sent by Email**	3	2	3	4	2	35%
11	**Consent Given Sent by Email**	2	2	4	3	2	38%
12	**Login Authentication**	2	2	2	3	2	42%
13	**Confirmation Button by Email**	2	2	3	4	1	44%
14	**Submission of a Document**	2	2	3	2	1	52%

F = **F**reely given;
A = **A**uditable;
E = **E**xplicit;
U = **U**sability;
T = **T**rustworthy.

3.2.1 Agreement Analysis

In this subsection we evaluate the agreement among the raters regarding the different parameters in the various mechanisms.

Methods

Interobserver agreement among experts was assessed using the proportions of agreement (PA). R software ('obs. agree' package) was used to compute the PA [22]. Ninety-five percent confidence intervals (95%CI) were calculated [22]. We present the agreement among observers in Table 3.3. From the results we can conclude that the agreement among observers is very low, in fact when a given observer rates one method the probability of another one will give the same rate ranges from 30% to 40%. This shows that the risk classification is not consensual, further more given different experience

levels in this subject as in our study. To overcome the lack of agreement the median of raters was considered as the best approach.

Table 3.3: Agreement between observers on the mechanisms of consent.

	PA [95% CI]
Freely Given (F)	0.30 [0.25,0.38]
Auditability (A)	0.41 [0.35,0.49]
Explicitness (E)	0.31 [0.25,0.39]
Usability (U)	0.38 [0.31,0.45]
Trustworthy (T)	0.33 [0.29,0.40]

PA = Proportion of Agreement;
95% CI = 95% Confidence Interval.

3.2.2 Analysis of Risk Results

Analyzing the Table 3.2, we can conclude that the best methods to achieve a consent that can be used are the ones that require a valid signature that qualifies the identity of the data subjects, whether it is a normal signature, if the request for consent is personally requested, or an electronic signature, with a qualified electronic certificate, if the consent is given electronically. However, both methods fail in usability, with the signature that requires the presence of the data subject considered worse than the electronic one, since it geographically limits the data subject. Even though some measures are needed to optimize the *Paper Signed in Person*.

Nevertheless, they are not the only methods that can be applied, others with higher risk can also be used, but their risk must be considered in the DPIA. It should be noted that the choice of consent mechanism should be weighted according to factors such as context and data type used in the processing operation.

The digital signature with a certificate that is not validated by a competent identity, a signature with an advanced electronic certificate, it must be uniquely capable of identifying and linking its signatory and guarantees the integrity of the document. So, a qualified electronic signature is an advanced electronic signature with a digital certificate that has been encrypted by a secure signature creation device.

An *Email from Data Subject* might suffer from spoofing attacks [33], as well as man-in-the middle-attacks [31]. However, if the email is signed with a qualified certificate

then its security and reliability is the same level as the *qualitative electronic signature* method.

The limitation of the universe of individuals ends with the *Code sent by email*, but this method is, as well, susceptible to attacks of man in the middle [31] and spoofing attacks [33].

Code sent by SMS, like the previous, it does not create limitations on the size of the universe of individuals but might suffer from spoofing attacks as well as man-in-the-middle attacks. Although, this method is safer than the previous one since it uses different channels of communication.

The methods using a *Digital Mobile Key* identifies the real identity of the data subject, because they are related to a qualified electronic certificate.

Digital mobile key, since combines the "Two-factor Authentication by SMS" and "Digital Signature" have a mix of obstacles identified by both mechanisms, on the one hand it limits the universe of individuals, since the government must provide this method, and on the other hand, the code can be intercepted by a man-in-the-middle attack, but no longer suffers from spoofing attacks. Even so, this mechanism is more practical than the direct use of the smart card (Digital Signature), although with less guarantees.

Digital mobile key with smartphone application combines the "Two-factor Authentication by application" and "Digital Signature" has a limited universe of individuals - government must provide, as well, this method - and the data subject are required to have a smartphone. Its security level is also dependent on the quality of the application. Nevertheless, this mechanism is more practical than the direct use of the smart card (Digital Signature) and, if the application is well implemented, the risk of suffering man-in-the middle attacks is considered low.

The risk associated with *login authentication* is related to the authentication and identification mechanism, which, for example, can only be the pair username/password or can be by duo factor. The controller must also guarantee that until the data subject defines what he/she consents, his/her personal data cannot be processed. However, this method is not ideal because the data subject must already be associated with the system, without being informed of its functionalities and requirements.

A *code to access the consent given* and *consent given sent by email* might suffer from spoofing [33] and man in the middle attacks [31]. However, the second one is more explicit, since its needed an extra action from data subject to consent. On the other hand, the first allows the data subject to access the form whenever wanted.

A *voice call*, if correctly applied can be a reliable mechanism since it can promote the freely given factor. The recording of the call demonstrates the information provided to data subjects and their reaction to it. Nevertheless, the recording of voice can be very intrusive to the privacy of the data subject and generate new risks to the data subject, in addition, it is also subject to impersonation attacks and the information given to the data subject might be excessive to receive entirely by the phone (limiting the freely given factor) and lastly, the record can be modified.

A *submission of a document proving identity*, like the previous, is subject to spoofing and man in the middle attacks because a document is easily forged. This mechanism also gives a false sense of security since a printed document cannot guarantee the digital identity of a person.

The *confirmation button of the consent by email* in addition of being susceptible to both man in the middle [31] and spoofing attacks [33]. Besides, due to the easiness of masking malicious websites, attackers might use this method to perform phishing attacks. Therefore, this mechanism should not be used since it might become a threat to cybersecurity and data protection.

3.3 Cleaning Personal Data from Databases for Marketing Purposes

A special case to take in consideration during the period of adaptation to GDPR is the restructuring of the processing operations already in process, including the conformity with the consent requests.

The marketing campaigns will be mostly under two lawfulness of processing, the consent condition or the legitimate interest, after establishing a contract. If they follow by consent, then they have to verify if the consent has already been collected according with GDPR, if not the companies have to recollect or collect the consent of the data subjects already on their databases.

The request for consent must be granular, which means that the controllers must identify the marketing campaigns they have and let the data subjects select which ones they want to receive notifications.

In this case, the consent request is directed, since the controller just wants to validate the consent of the data subjects already under processing operations, and consequently,

already has a channel to communicate, with the email being the most common. To be noted that, using a channel to communicate with the data subject without being initially consented is against GDPR, so this request must be done before the application of GDPR.

So, the easiest way to request the consent would be by the method already described, namely *Confirmation Button by Email*, but this mechanism is not a good option since phishing attacks can be easily masked using this approach. An SMS with a link is also considered insecure, since it automatically directs to a webpage that is susceptible to phishing attacks.

These methods can be accomplished in two different ways. The first is by sending an email for each purpose, resulting in the data subject with a full message box and ignoring them all. The alternative approach is to send a link where all the purposes are listed.

If the email sent to data subjects does not have a link but asks to go to the company's webpage in order to select their preferences, then there is no exposure to phishing, but the adhesion of the data subject is lower, since it requires more actions from them. As with the previous, it has the level of risk as sending an SMS.

Likewise, if the data subjects are requested to answer to the email identifying the purposes they consent, then the email sent by the company should identify the purposes of the processing operations.

Another way frequently used by the marketing companies, is the *Voice Calls*. This mechanism is not vulnerable to phishing attacks but is geographically and time limited to the call centers and employees of the company with the aggravating factor that is the most intrusive mechanism presented.

If the company has a webpage where authentication is needed, then the consent can be requested in the portal, by the *login authentication*. This consent method is the least intrusive for data subject, since only the interested data subjects access to the portal.

In the consent request, companies should be careful not to overload the data subjects as it may lead the data subjects to ignore the requests. Thus, the company can request consent only for campaigns that the data subject is already subscribed to. And at a later stage, when the data subject asks to change the settings, the company might present all available campaigns.

Chapter 4

Proposal of DPIAs for Information Systems

The GDPR sets out different obligations depending on the nature of the personal data that is processed, as well as the ways to collect consent or the level of security measures to be implemented.

In general, information security can be defined as the preservation of the confidentiality, integrity and availability of information. Thus, a set of controls, such as policies, practices, procedures, organizational structures and software functionalities shall be implemented.

4.1 Problem Found: Motivation

During the past years, we have witnessed massive record data breaches because of many information security incidents involving PIIs that have affected both individuals and organizations, some of the incidents involving legal liability, identity theft, and recovery costs.

With the introduction of the GDPR, the organizations (controllers and processors) became responsible for the compliance of the regulation without the comfort of the DPAs' opinion, by the principle of accountability, and non-compliance with these obligations may result in significant financial impacts, with fines with values never seen, accompanied by reputation depreciation, since data breaches must be disclosed to data subjects if breach entails risks to individuals

This combination originates that the non-compliance is no longer suitable for the organizations, forcing them seek safe and privacy-friendly alternatives for processing personal data, and adjust their procedures to comply with the GDPR requirements. Thus, the organizations tend to choose procedures and tools already tested and accepted for the processing of personal data, to protect their privacy networks and PII, to align with the increased usage of information and communication technologies.

Therefore, GDPR compliant seals have been of big relevance, since, in one hand we have the organizations that process personal data, that try to be as prepared as they can for this regulation, and with that, they look for solutions that ensures compatibility with the regulation. On the other hand, we have software suppliers that see this regulation as a business opportunity and bet on systems that guarantee to be GDPR compliant.

However, since the GDPR tunes the way the data is processed, and it is independent of the support or the degree of automation of the system, we do not believe that privacy seals are the proper way to show systems GDPR compliant. Nevertheless, as the ISs are very relevant in the elaboration of DPIAs, and the right implemented measures can minimize the risk of impact on personal data, this is also referenced in the GDPR, where defines that technological measures should be applied in the systems, like the data encryption and the recording of the access logs[Art. 32.o] [19].

Thus, since the ISs that support the processing must also have a bearing on the risks involved in data processing and its impact on the privacy of individuals, the software must be secure. Consequently, these, too, must undergo a process of risk analysis and privacy impact assessment, which will then be weighed along with the privacy impact assessments of the treatments. So, combining privacy impact assessment of the data processing with the privacy impact assessment of the systems that are used during the processing, we obtain a grade of how compatible we are with GDPR.

4.2 Proposal of DPIA for Information Systems

In order to prepare the evaluation proposal, a survey of the questionnaires already proposed was carried out, being, in this case, the questionnaire from Microsoft [34], the Spanish Data Protection Authority (AGPD - *Agencia Española de Protección de Datos*) [8] and the Frensh Data Protection Authority (CNIL - *Commission Nationale de l'Informatique et des Libertés*) [16]. Finally, to ensure that the evaluation covers the essential points of a secure system, a mapping of GDPR with ISO 27001 was, also,

carried out.

It is important to notice that the risk and the impact of the failure of an IS is totally related with the kind of personal information that the IS processes. Therefore, in order to assess the risk, we must (i) identify the potential effect (illegitimate access to data; unwanted modification of data; temporary or definitive unavailability of data); (ii) identify the sources risk; (iii) identify the possible threats; (iv) determine the existing or planed measures; and (v) assess the severity and likelihood.

4.2.1 Mapping of ISO 27001 in GDPR

Privacy has largely been a matter of law and policy, traditionally working along a spectrum that's context dependent, while security has largely been a matter of technology and policy, traditionally working in binary states. However, with GDPR, "adequate security" is now mandatory by law. With these complex pieces of legislation, there is a emerging class of technologies to help privacy teams understand and comply with them operationally [23, 39].

Increasingly, this means that more than ever the security of technologies supports privacy, with the idea of *data protection* [23, 39].

GDPR focuses specifically on the protecting and appropriately managing personal data. ISO 27001 focuses more broadly on creating an information security management system (ISMS) to prevent data loss or ex-filtration and ensure that a institution's information security posture can be maintained, and incidents identified, logged and reported. This includes guidance on how to handle and protect personal data in a secure, trustworthy manner [23, 39].

In Article 32, the GDPR states that organizations "*... shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk...*" [19], it also mandates other security-related points [23, 39].

For the elaboration of the semiautomatic privacy assessment, a mapping of ISO 27001 with GDPR was made, identifying which points of the ISO 27001 were referenced in GDPR. At first sight, we can notice that here are eight specific areas where ISO 27001 directly supports GDPR compliance, as demonstrated in the table 4.1 [23, 26, 39].

Table 4.1: General areas directly supported in GDPR and ISO 27001.

Areas directly supported	**GDPR**	**ISO 27001**
Management of personal data	GDPR "...lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data". [Art. 1]	ISO 27001 mandates the organizations "to ensure that information receives an appropriate level of protection in accordance with the importance to the organization", so personal data should be considered as important to the organization. [A.8.3]
Third-party risk management (TPRM)	GDPR stipulates that businesses that "...processing by a processor shall be governed by a contract or other legal act...". [Art. 28]	ISO 27001 mandates the organizations "to ensure protection of the organization's assets that is accessible by suppliers". ISO mandates the organizations "to maintain an agreed level of information security and service delivery in line with suppliers' agreements". [A.15.1, 2]
A documented process for regularly assessing the effectiveness of security controls	GDPR mandates that controllers to maintain a record of processing activities under its responsibility. GDPR mandates to create "...a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing". [Arts. 30, 32]	ISO 27001 mandates the organizations to have the processing documented, with the implemented controls. Those documents should, after, be updated and audited. [7.5]

Continued on next page

Table 4.1 – continued from previous page

Areas directly supported	**GDPR**	**ISO 27001**
Availability, integrity and confidentiality of data processing systems	GDPR mandates that the organizations must implement the appropriate technical and organizational measures to ensure a level of security appropriate "...to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services". [Art. 32]	ISO 27001 mandates the organizations "to ensure that information security is an integral part of the ISs...". [A.14.1]
Risk assessment	GDPR mandates that organizations conduct risk assessments to ensure they've identified major risks to EU citizens' personal data. [Art. 32]	ISO 27001 mandates the organizations "to preform information security risk assessments". [8.2]
Data encryption	GDPR advises the organizations to implement measures to mitigate the risks, such as encryption. [Art. 32]	ISO 27001 mandates the organizations "to ensure proper and effective use of cryptography to protect..." the CIA of information. [A.10.1]
The ability to restore access to personal data in a timely manner	GDPR mandates that organizations should have "...the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident". [Art. 32]	ISO 27001 mandates the organizations "to protect against loss of data". [A.12.3]
Breach notification	GDPR mandates that organization must notify authorities within 72 hours of when a breach involving personal data is discovered. If the risk to them is sufficient then data subjects should also be notified. [Arts. 33, 34]	ISO 27001 mandates the organizations "to ensure a consistent and effective approach to the management of information security incidents, including communication on security events and weakness". [A.16.1]

A more thorough analysis of the ISO mapping in GDPR can be found at Mapping ISO 23445 in GDPR or in the next url: `http://www.alunos.dcc.fc.up.pt/~up201206845/Mapping.pdf`, where an analysis is made of which controls of ISO 27001 are applied to each paragraph of the GDPR. However, the controls A.18.1.4 and A.18.2.2 of ISO 27001 are not considered as they define that applicable regulations and policies implemented should be followed, and the Clause 4 for the fact that organizations must guarantee that the scope encloses all phases of the treatment of the personal data, which applies for the entire Regulation.

4.2.2 DPIAs Structure

The Information Systems DPIA is composed by two risk assessment questionnaires. On the one hand, there is a form assessing how the organization is technically secured and privacy friendly, like workstations or physical security. On the other hand, the second questionnaire assesses how a specific IS is technically secured and privacy friendly, like the configuration of the server or website, the software development or how the data processors are managed. The rest of the sections bellow identified applies to both of surveys.

Consequently, with the inputs given by CNIL, AEPD and ISO 27001, we concluded that in order to consistently assess an IS on the level of security and data protection, the following points should be included in the assessment:

Identification and authentication
This group assesses how the users are created and managed, in order to ensure that the user only access the data that he/she needs, he/she must be associated with a unique identifier and must authenticate himself/herself before any access to personal data.

Access management
In this group, we assess how the users created in the previous group are managed, if the access is limited to only allow access to data that the user really needs.

Access logs and data breaches management
In this section, the access log and management of incident procedures allowing to react in the event of data breach (breach of confidentiality, integrity or availability)

are assessed. These mechanisms help identify fraudulent access or abusive use of personal data, or to determine the source of an incident. Thus, incident recording and management measures must be implemented, recording relevant logs, and ensuring that these logs cannot be changed.

Workstation
This group assess how well the organization is prepared to prevent fraudulent access, virus execution, or remote control, since the risks of ISs intrusion are significant and workstations are key points of entry.

Mobile data processing
With the increasing use of laptops, pen drives and smartphones, it becomes necessary to be prepared for data breach after theft or loss of such equipment. With this group, it is intended to anticipate data breach after the theft or loss of a mobile device, evaluating the measures implemented to ensure the CIA of the personal data.

Internal network
This group evaluates the security measures implemented in the internal network, such as what is allowed in the internal network, since only the network functions necessary for the implemented processing must be authorized.

Servers
Server security should be a priority because they centralize a large amount of data. Thus, this group strengthens the security measures applied to the servers, in order to protect personal data.

Website
This group aims to ensure that the best basic practices are applied to sites, for example, each site must guarantee the CIA the information it sends or collects.

Business continuity
A business continuity plan that anticipates possible incidents (for example, hardware failure) should be prepared. Thus, this group evaluates how regular backups are performed to reduce the effect of undesired data loss and how and when backup copies

should be made and tested.

Archive

Archived data are those that are no longer used periodically but have not yet reached the end of the data retention period, for example because they are held for use in case of litigation. Thus, this group aims to evaluate how files are protected, especially if archived data is special data that can have serious impacts on data subjects.

Supervising maintenance and data destruction

This group evaluates the measures implemented to ensure data security at all times in the hardware and software life cycle. Thus, maintenance operations must be supervised, such as data access control by service providers. And, at the end of the contract with the service providers, the data should be deleted before discarding the hardware [14].

Data processors management

The GDPR argues that the data processing should be supervised by the controller, including the security measures implemented. Thus, personal data communicated or managed by subcontractors must be processed with security guarantees, which must be defined by the controller. This group evaluates how the organization is ensuring GDPR compliance by its processors [14].

Shares and transfers

This group aims to assess the security of all transfers and sharing of personal data. Evaluating, for example, what medium is used to transmit personal data (such as electronic mail services that are not a secure means of communication without additional measures - message encryption). A simple manipulation error can result in the disclosure of personal data to unauthorized recipients and therefore interferes with people's right to privacy. In addition, any entity with access to the messaging servers in question (especially the senders and recipients) may have access to its content [14].

Physical security

This group aims to assess the security of facilities hosting IT servers and network equipment, which must be protected so that access to facilities is controlled to prevent or slow down unauthorized access. These measures should be applied to paper files or IT equipment, especially servers [14].

Software development

This group is based on privacy by design and privacy by default. Thus, it assesses if the security and privacy was integrated as soon as possible in the projects. Defending that privacy must be integrated with software development from the different stages of the project in order to give data subjects better control over their data and to limit errors, losses, unauthorized modifications or misuse of personal data in applications [14].

Confidentiality, integrity and quality of data

Two of the key concepts of information security are the confidentiality and integrity assurance of information. Thus, measures should be defined for the confidentiality and integrity to be safeguarded, such as hash functions to ensure data integrity, digital signatures, which, in addition to ensuring integrity, are also capable of verifying the origin of the information and its authenticity, and finally, the encryption that allows to guarantee the confidentiality of a message. Therefore, in this group we evaluate the mechanisms implemented to guarantee this concept [14].

Audit and responsibility

One of the major obstacles to data security and privacy is how users treat and apply the defined procedures. Thus, in this group we evaluate how the organization makes users aware of the privacy and security challenges of the organization and how internal policies are transmitted [14].

Regarding the structure of the questionnaire, all questions are elaborated in such a way that positive responses benefit (adding the value of the question) and negative responses are not valued, increasing the risk. In the check boxes the worst-case scenario is calculated, with the risk representing the worst case.

Since the DPO does not necessarily need to have technical skills, our estimate is that most DPOs will have a legal background, so the questionnaire attempts to assist the DPO in its GDPR compliance task, thus for each negative answer a suggestion for improvement is made available, which is part of the final risk analysis report. In this way the DPO does not need to know the technical characteristics of the ISs, it only needs to ensure that the corrective measures are part of an action plan to be carried out.

4.3 Classification Methodology

The DPIA template has in total 221 questions, which are given an importance scale rating of 4 (1: Not important to be present - 4: Essential to be present). However, these questions are subdivided in two assessment questionnaires, one that is applied to each IS, with 133 questions, and other applied to the policies and common procedures within the organization, with the rest of the 89 questions.

Thus, to carry out the risk analysis, seven raters were asked to evaluate the questions according to their experience the different parameters in the various mechanisms. For the selection of raters, their experience and diversity in the scope of activities were taken into account, and then selected according to convenience. Thus, the raters experience are: one MSc student, two PhD student, one Assistant Professor with more than 10 years' experience, one Associate Professor with more than 20 years' experience, one security expert with more than 10 years' experience, and one system administrator with more than 10 years' experience.

4.3.1 Personal Data as Variable

It should be noted that the level of risk is also calculated taking into account the data that the IS supports, so if the IS contains sensitive data the risk generated in the privacy impact assessment is higher than in a system that contains only common data. Thus, the assessment contains a pre-definition of the IS that qualifies the IS between three levels according with the information held in the IS. Although, GDPR only distinguish between special data and non-special data, we believe, like AEPD, that there should be an intermediate level of data protection, as some data such as administrative and criminal convictions and infractions, capital and credit solvency, location or movements, may not be sensitive information, but combined with other data, or even if the data results in something specific, can become discriminatory. Thus, the following segregation is proposed:

Low level
All ISs should adopt a minimal security measures, in this level all the personal data is enclosed.

Middle level
To the ISs that processes data that might become discriminatory, but usually is not (like data related to administrative or criminal infractions, equity and credit solvency, location or movements) should be required to adopt more security and privacy friendly measures.

High Level
Finally, for the ISs that processes special data (racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation) and the data that gives rise to a Data Protection Impact Assessment (DPIA) (video surveillance in public places and profiling) the security and privacy requirements should be even more demanding.

4.3.2 Agreement Analysis

In this subsection we evaluate the agreement among the raters regarding importance of the questions to ensure the privacy of personal data in information systems.

Methods
Interobserver agreement among experts was assessed using the proportions of agreement (PA) [22]. R software ('obs. agree' package) was, again, used to compute the PA. Ninety-five percent confidence intervals (95%CI) were calculated [22]. We present the agreement among observers in Table 4.2. From the results we can conclude that the agreement among observers is very low. This shows that the definition of the most important tools for guaranteeing data protection is not consensual. To overcome the lack of agreement the median of raters was considered as the best approach to define the value in the questions on the SIs' DPIA. In relation to observers, there were three observers, who among them had an agreement above 50%, varying between 50% and 66%, which are those that work more directly with data protection, which shows that although the agreement is still not ideal, the valuation of quotations by observers is very dependent on their contact with this matter and that this type of risk analysis should be studied and standardized.

In addition, it has also been found that if the scale only differentiate the important of not important, has a scale of 2 instead of 4, than the agreement highers significantly,

as showed in Table 4.2.

Table 4.2: Agreement between observers on the importance of questions SIs' DPIA.

	PA [95% CI]
Scale with 4 levels	0.39 [0.37,0.41]
Scale with 2 levels	0.71 [0.69,0.74]

PA = Proportion of Agreement;
95% CI = 95% Confidence Interval

A study of interobserver agreement among experts was, also, assessed per section, where a lack of agreement is uniform between all sections. We present the agreement among observers in Table 4.3.

Table 4.3: Agreement between observers on the importance of questions SIs' DPIA per section.

	PA [95% CI]
Identification and authentication	0.35 [0.30,0.42]
Access management	0.43 [0.35,0.52]
Access logs and data breaches management	0.36 [0.31,0.42]
Workstation	0.38 [0.31,0.46]
Mobile data processing	0.34 [0.28,0.41]
Internal network	0.42 [0.33,0.54]
Servers	0.45 [0.37,0.54]
Website	0.32 [0.26,0.43]
Business continuity	0.36 [0.32,0.51]
Archive	0.36 [0.27,0.50]
Supervising maintenance and data destruction	0.44 [0.36,0.53]
Data processors management	0.33 [0.27,0.37]
Shares and transfers	0.43 [0.30,0.56]
Physical security	0.38 [0.31,0.43]
Software development	0.40 [0.32,0.45]
Confidentiality, integrity and quality of data	0.39 [0.30,0.50]
Audit and responsibility	0.38 [0.33,0.43]

PA = Proportion of Agreement;
95% CI = 95% Confidence Interval

Chapter 5

ISs' DPIA' Evaluation in Organizational Environment

5.1 Implementation

In order to, initially, assess our proposal an environment for testing was implemented. Thus, a virtual machine was configured in order to install the LimeSurvey platform [21], so that as the DPIA was answered by the organizations, they received automatic feedback on the state of the IS and its recommendations.

The LimeSurvey platform offered, direct or indirectly, all the tools needed to prepare the questionnaire, such as multiple-choice questions, yes/no, check-boxes. And allowing the quotation of the questions from the logical ability of the questions, printing recommendations from the ability to text display, as well as drawing risk graphs by inserting html code in the same type of question (text display).

However, for the elaboration of the DPIA as intended, the tool is not the most facilitating, since after the questionnaire was accessible to fill, it does not allow changing the structure of the questionnaire, adding options in check-boxes, removing or adding questions. It is necessary to respond from the beginning to the questionnaire whenever there is a need to make one of these changes. Also, the system should allow different access levels like LimeSurvey platform does not have.

Currently, the ISs' DPIA can be found on the TekPrivacy platform, which no longer has the limitations found on the previous platform, which is still a work in progress, but it is a platform specially though to the specificity of GDPR.

5.2 Applying ISs' DPIA in Organizational Environment

After defining of the information systems DPIA, it was important to evaluate it in a real-world scenario and to evaluate its compliance with the GDPR. The following institutions were studied: MP, GO Porto and APDL. This IS analysis allowed the organizations to perform a more accurate risk analysis of the processing and to learn where its efforts should be applied. And, on the other hand, it allowed us to improve our risk analysis by identifying the biggest obstacles in the current market.

Although the organizations considered are public, that may cause some bias, there are quite different in dimension, one of the organizations having 3 ISs, other 26 and the last one 73.

In the table 5.1 we present some properties used to assess for the ISs for each organization, note that, not all systems have the properties indicated below. When the number of ISs with the property implemented is very similar with the number of ISs without it, then it is considered as partially accomplished, Finally, might appear the Not Applicable option, when the organization considers that the property is not relevant to the ISs. The organizations in the table are not identified and were randomly added to the table.

Table 5.1: Analysis of the mechanisms implemented in the information systems of the organizations under study.

Question	**O1**	**O2**	**O3**
There is a procedure for allocating, distributing, and storing passwords	✓	✓	✓
Passwords validator prohibits the reuse of passwords	✓	✘	✗
Authentication mechanisms are working correctly	✗	✓	✓
Strong authentication mechanism is used	✗	✗	✗
Appropriate measures for task delegations are created	✗	✓	✗
Session data is validated	✗	✘	✓
The session closes after some time of inactivity	✓	✗	✓
Logout functions have been audited	✓	✗	✓
The session generation method was audited	✓	✗	✗
There is no possibility to deactivate the logs	✓	-	✗

Continued on next page

Table 5.1 – continued from previous page

Question	**O1**	**O2**	**O3**
Monitoring is performed on access controls of third parties	✗	✗	✗
Personal data stored on off-site treatment is expressly authorized	✓	✓	-
The level of security corresponds to the type of system treated	✓	-	-
The relationships between the BDs are performed by internal identifiers	✓	✓	✓
Backup procedures are reflected in the security policy	✓	✓	-
The backups ensure the reconstruction of the data to the state they were in before the loss or destruction occurred	✓	✓	-
Measures are taken to prevent retrieval of personal data discarded	✓	✓	-
Discarded data or documents are removed from inventory	✓	✗	-
Temporary files/copies meet the corresponding security level	✗	-	✓
Temporary files were destroyed/deleted when they were no longer needed	✓	-	✗
The transfer of data is made safely (guarantees CIA)	✗	✘	-
Data or documents entry and exit is recorded	✗	✗	✗
The code is peer reviewed	✗	✗	✗
Break the Glass mechanisms are implemented	✗	✗	✗
Pseudonymization and/or anonymization algorithms are implemented	✗	✗	✗
The information contained in the data or document is classified	✗	✓	✗
There is an updated inventory of the data / fields stored	✓	✘	✓
Some personal data is encrypted	✓	✗	✗
Cryptographic keys are changed frequently	✓	-	✗
Audit activity encompasses all servers and layers	✓	✓	-
The corrective measures proposed by previous audits were implemented	✓	-	-
The corrective measures of the audits were effective	✓	-	-
Internal vulnerability and threat tests are scheduled	✓	✗	-

O1 = Organization 1;
O2 = Organization 2;
O3 = Organization 3;

✓= Accomplished;
✘= Partially accomplished;
✗= Not Accomplished
- = Not Applicable

With this survey we observed that, although the organizations already considered the security of their ISs, they did not have in mind privacy issues. Also, the procedures are defined and are being applied, but most of them are not documented. Still in the scope of the policies, we also observed that although some policies are defined and documented, some of the ISs do not meet the defined requirements, in part because

these characteristics should be implemented from the supplier and were not considered by design.

Most systems only use password as an authentication mechanism, which is a bad practice, due to repetitions, reuses and standardization of passwords. And the organizations do not control the access of third parties (service providers) during the monitoring of the ISs.

The sections with lower risk, so the sections more compliant with GDPR, are the *Business continuity*, *Identification and authentication* and *Physical security*. The first one, because most organizations try to be prepared in case of incident and are already structured to perform good backup procedures. The *Identification and authentication* has been discussed for some time, even though it is constantly improving, organizations are aware of the best practices of identification and authentication. Finally, *Physical security* is a subject studied for some time with well-defined procedures of conditions of the data centers and access control to these rooms.

On the other hand, the sections *Audit and responsibility*, *Software development* and *Confidentiality, integrity and quality of data* are below the expected, because are they are more related to privacy and security, a new concern of the organizations. In first place, we have auditability and responsibility that, although ISs have access logs, they are not verified and are not audited in order to guarantee their quality and relevance, nor is it common practice to test for vulnerabilities and threats to ISs. Another major problem encountered is the development of software, which although is very advanced, it did not consider concepts such as privacy by design and privacy by default. Lastly, *Confidentiality, integrity and quality of data*, to which the usability-security relationship is very important, it is necessary to carry out a study on what measures to implement to guarantee confidentiality and integrity, for example, under what conditions the data must be encrypted and signed, once whereas these measures have a very high impact on the organization's.

5.3 Assessment of the Proposed DPIA

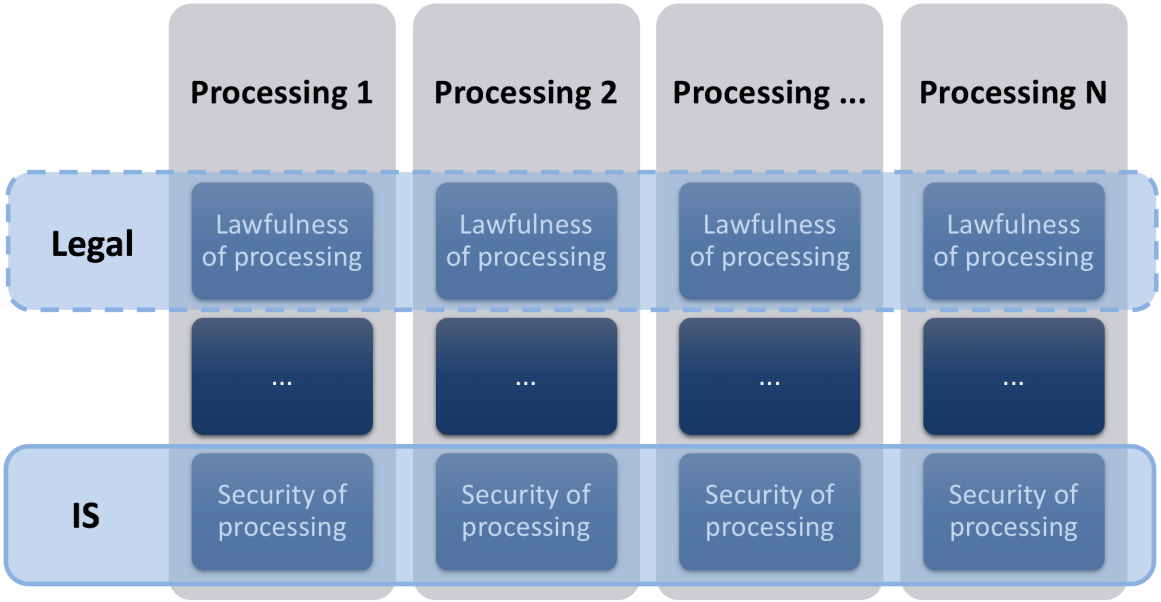
The record of processing activities, defined in GDPR at Article 30, must document how the processing activities is lead. Therefore, the security measures undertaken and the ISs that gives support to the processing is important in the risk analyses and, if necessary, in DPIAs. Given that several processing activities can share the same IS, we

concluded that the IS should be assessed independently of the DPIA. By doing this, one can evaluate the risk of each processing activity and just add the results for the supporting IS.

Furthermore, the isolation of the risk assessment of the IS, also permits the organization to identify which IS needs more attention, because supports more processing activities, or the processing activities that the IS supports are sensible, or because is the IS with less security measures or downgraded.

Besides, the ISs' DPIA considering privacy and security issues of the system and is completed by the technical team, leaving to the responsible the characterization on how the processing is made and how the data subjects can execute their rights. This aspect, also, made us think that, maybe, the legal aspects should also have an isolated risk assessment, like exemplified in figure 5.3, but this assessment has not yet been developed.

Figure 5.1: Processing characterization and compliance assessment



One of the main difficulties encountered in the elaboration of the assessment questionnaire was the definition of the number of questions, since it becomes a trade-off between usability and utility. For the evaluation be consistent, it is necessary to have a significant number of questions. On the other hand, if the assessment has many questions, then it becomes tiresome and of low usability.

The problem of the DPIA per IS with many questions was made worse by the fact that

the organization needed to answer to a questionnaire by IS and some of the answers to the questions were the same among the ISs.

In order to reduce the number of questions to be answered for each IS, the original questionnaire was divided in two, depending on whether the question was related to the organization's security and privacy principles (independent of the IS), or whether the question was related to a specificity or characteristic of the IS.

Another problem encountered in risk analysis is the diversity of application. Since there are several ISs with different requirements and different applications, it was necessary to make the questions abstract enough so that each system specification could be overcome by the questions, but on the other hand it would allow us to assess in a specific way the criterion that we searched for, being that in some questions the only solution found was to give the permission of "not applicable". The fact that we open the door to the "not applicable" raised another problem, since professionals often, when the answer is negative tended to classify it as "not applicable".

However, since this evaluation must be adapted to the reality of the organizations, this is a work in progress model, which is being improved as it is applied in practice, that is, organizations are completing the risk assessment for their ISs. This improvement may involve, for example, the adjustment of the questions, as well as the re-evaluation of the evaluation of the questions.

Chapter 6

Conclusions and Future Work

As in all risk evaluation, risk analysis for the protection of personal data must have a balance between functionality (complies with GDPR requirements), usability (easy and quick to use) and security (do what is expected and nothing more). Thus, all procedures, methods, mechanisms and tools chosen for data processing must be evaluated and their risk should be weighed, so a trade-off from where the organization wants to prioritize must be performed. In this work, we aimed to facilitate the risk analysis and its evaluation for consent acquisition and for the ISs that support the processing.

6.1 Research Summary

As noticed, the mechanisms with greater levels of security are those that have a valid signature according to the environment of consent. If printed paper is used, then a normal signature guarantees the identity of the person. In the case of a digital reality, the signature needs to be digital, using a qualified electronic certificate. So, the only mechanisms identified that are considered as trusted and guarantees a level of compliance as defined by GDPR are the ones that most minimize the universe of data subjects, thus limiting the scope of processing operations. Contrariwise, the methods *confirmation button of the consent by email* and *submission of a document proving identity* are the ones with lower levels of trustworthiness. Nevertheless, the mechanism that requires less actions from data subjects, the *Confirmation Button by email*, is, as expected, the one with higher usability. But this easiness, causes a lack of confidence in the mechanism. In the opposite case, the method with lower usability is the *Paper Signed in Person*.

Some common weaknesses and strengths were found in the organizations evaluated. As indicated in the previous chapter, the sections best prepared to ensure the security and protection of personal data are *Business continuity*, *Identification and authentication* and *Physical security*, mainly because they are part of the principles to be taken into account for the continuity and execution of the business, and for that reason already well-known and practiced care. On the other hand, the less prepared sections are *Audit and responsibility*, *Software development* and *Confidentiality, integrity and quality of data*, which in turn ensure security, privacy and data protection that is a new concern of organizations due to the computer attacks that have occurred and due to GDPR. In any case, few systems are prepared for these factors and organizations with procedures that guarantee security and data protection.

Similar to the idea of the AEPD - Evaluates, the ISs' DPIA attempts to assess data security and privacy conditions, but only those that are stored in ISs, since this semiautomatic risk analysis evaluates ISs privacy issues and not information security. However, unlike the CNIL platform, this evaluation is more closed, trying to facilitate the work to the controllers, evaluating the risk from the yes / no answers. It also tries to segregate the evaluation areas so that each analysis is performed by a technician and the common processes, so they do not need to be evaluated for each treatment, although it is still a work in progress.

6.2 Current Limitations

As in several other parameters in the risk analysis, the consent method implemented should also depend on the sensitivity of the personal data involved in the processing operations. There is no single universal accepted method, as the trade-off between the several parameters introduced in the GDPR is quite diverse. Notice that the GDPR reinforces the risk analysis and that the institutions should measure the risk and either mitigate it or accept it. We stress that in an online environment the notion of freely given is really hard to assure as we are not able to see the consent expression, this is similar to the online voting where the freely given is also a key aspect [30].

In the study of the definition of an ISs' DPIA, we found that privacy and protection of personal data are novel, since most systems in production did not take into account privacy by design and privacy by default. Thus, most ISs are not prepared to ensure the protection of personal data, and the measures that can be taken at this time are of great importance to organizations, both procedural and financially, since it would

be necessary to implement new tools in the system and this correction can lead to a decrease in productivity. Moreover, the ISs' DPIA was only tested in 3 organizations, with the 3 organizations being public, so it is necessary that the risk analysis be evaluated in a larger and more varied sample of organizations.

Other limitation of the analysis is the low agreement between observers, demonstrating that there should be more debated and consensual issue, in order to facilitate organizations when choosing the best practices to ensure compliance with the GDPR.

6.3 Future Work

As future work we have two different lines of research, these being at the level of the request for consent, and the isolated risk analyzes.

So, for future work, we plan to evaluate the agreement between the raters by using proportions of agreement. This will be particular useful to validate the risk rating that we are proposing. The risk rating, which was statistically validated in this work, will hopefully play a crucial role in the Data Protection Impact Assessments. Regarding consent, we believe that it is urgent to develop a usable zero-knowledge protocol to fulfill the data minimization requirement.

For future work of ISs' DPIA we plan to continually improve as we have more IS to analyze and study the reality of organizations. In addition, we also intend to create more segregated risk analyzes, so we need to check what common points have all the processing and when does this segregation benefit the organization in terms of availability of costs and resources. As previously indicated, for this work, we are already thinking of separating the legal contents of the remaining analysis, since, it has to allocate a specialized professional to carry out this analysis.

6.4 Conclusion

In conclusion, although the regulation has been in place since May 25, 2018, the GDPR adaptation activity still requires substantial additional work. Security and privacy are far below the technologies currently used. Thus, there are neither mechanisms for collecting consent that fulfills the requirements imposed by regulators in a usable way, nor the ISs that handle the information does not have the protection and security

mechanisms expected. The problem goes beyond technical failures when there is no harmony on the impact that each consent mechanism has on a person's privacy and on the degree of importance of the characteristics in ISs. This is even more evident in today's processing model of personal data being used.

References

[1] Article 29 working party. https://epic.org/privacy/art29wp/. Last accessed on 2018-04-02.

[2] Universal declaration of human rights. http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf, December 1948. Last accessed on 2018-05-31.

[3] European convention on human rights. https://www.echr.coe.int/Documents/Convention_ENG.pdf, November 1950. Last accessed on 2018-05-31.

[4] International covenant on civil and political rights. http://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx, December 1966. Last accessed on 2018-05-31.

[5] Act 67/98 of 26 october. act on the protection of personal data (transposing into the portuguese legal system directive 95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data). *Portuguese Parliament*, October 1998.

[6] Constituição da república portuguesa. http://www.parlamento.pt/Legislacao/Paginas/\ConstituicaoRepublica\Portuguesa.aspx, April 2005. Last accessed on 2018-12-11.

[7] AEPD. Agencia española de protección de datos. https://www.aepd.es. Last accessed on 2018-06-17.

[8] AEPD. Evalua - test de cumplimiento de medidas de seguridad. http://www.servicios.agpd.es/Evalua/home.seam. Last accessed on 2017-12-10.

[9] Jens Bender, Özgür Dagdelen, Marc Fischlin, and Dennis Kügler. Domain-specific pseudonymous signatures for the german identity card. In Dieter Gollmann and Felix C. Freiling, editors, *Information Security*, pages 104–119, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[10] EU GDPR Portal: Powered by Trunomi. Gdpr key changes. `http://www.eugdpr.org/key-changes.html`, note = Last accessed on 2017-09-29. Last accessed on 2017-09-29.

[11] Ana Carvalho, Rolando Martins, and Luís Antunes. How-to express explicit and auditable consent. In *2018 16th Annual Conference on Privacy, Security and Trust (PST)*. IEEE, 2018.

[12] CNIL. Commission nationale de l'informatique et des libertés. `https://www.cnil.fr`. Last accessed on 2018-06-17.

[13] CNIL. The open source pia software helps to carry out data protection impact assesment. `https://www.cnil.fr/en/open-source-pia-software-helps-carry-out-data-protection-impact-assesment`. Last accessed on 2018-06-17.

[14] CNIL. Security of personal data. `https://www.cnil.fr/sites/default/files/atoms/files/cnil_guide_securite_personnelle_gb_web.pdf`. Last accessed on 2018-06-11.

[15] Giuseppe D'Acquisto, Josep Domingo-Ferrer, Panayiotis Kikiras, Vicenç Torra, Yves-Alexandre de Montjoye, and Athena Bourka. Privacy by design in big data: An overview of privacy enhancing technologies in the era of big data analytics. December 2015.

[16] Commission Nationale de l'Informatique et des Libertés. A new guide regarding security of personal data. `https://www.cnil.fr/en/new-guide-regarding-security-personal-data`. Last accessed on 2018-05-01.

[17] DocuSign. Understanding digital signatures. `https://www.docusign.com/how-it-works/electronic-signature/digital-signature/digital-signature-faq`. Last accessed on 2017-11-06.

[18] Directive 95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. `https://eur-lex.europa.`

eu/legal-content/EN/TXT/?uri=CELEX:31995L0046, November 1995. Last accessed on 2018-05-15.

[19] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC, May 2016. Last accessed on 2018-05-15.

[20] Ruth R. Faden and Tom L. Beauchamp. *A History and Theory of Informed Consent*. Oxford University Press, 1986.

[21] LimeSurvey GmbH. *LimeSurvey: An Open Source survey tool*. LimeSurvey GmbH, Hamburg, Germany.

[22] Teresa Henriques, Luís Antunes, and Cristina Costa-Santos. *An R package to assess agreement between observers*, 2013. The package includes two functions for measuring agreement. Raw Agreement Indices (RAI) to categorical data and Information-Based Measure of Disagreement (IBMD) to continuous data. It can be used for multiple raters and multiple readings cases.

[23] IAPP. Iapp-onetrust research: Bridging iso 27001 to gdpr. https://iapp.org/resources/article/iapp-onetrust-research-bridging-iso-27001-to-gdpr/. Last accessed on 2018-06-06.

[24] ICO. Conducting privacy impact assessments code of practice. https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf. Last accessed on 2018-01-19.

[25] ICO. Information commissioner's office. https://ico.org.uk. Last accessed on 2018-06-17.

[26] ISO/IEC 27001:2013(E): Information technology – Security techniques – Information security management systems – Requirements. Standard, International Organization for Standardization, Geneva, CH, October 2013.

[27] ISO/IEC 27018:2014(E): Information technology – Security techniques – Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors. Standard, International Organization for Standardization, Geneva, CH, August 2014.

[28] ISO/IEC 29100:2011(E): Information technology – Security techniques – Guidelines for privacy impact assessment. Standard, International Organization for Standardization, Geneva, CH, December 2011.

[29] ISO/IEC 29134:2017(E): Information technology – Security techniques – Guidelines for privacy impact assessment. Standard, International Organization for Standardization, Geneva, CH, June 2017.

[30] Chris Karlof, Naveen Sastry, and David Wagner. Cryptographic voting protocols: A systems perspective. In *Proceedings of the 14th Conference on USENIX Security Symposium - Volume 14*, SSYM'05, pages 3–3, Berkeley, CA, USA, 2005. USENIX Association.

[31] Seema Khanna and Harish Chaudhry. Anatomy of compromising email accounts. In *2012 IEEE International Conference on Information and Automation*. IEEE, jun 2012.

[32] E. Kosta. *Consent in European Data Protection Law*. Nijhoff Studies in European Union Law. Brill, 2013.

[33] Pandove Kunal, Jindal Amandeep, and Kumar Rajinder. Email spoofing. In *International Journal of Computer Applications*, volume 5, 08 2010.

[34] Microsoft. Microsoft benchmark tool: Detailed gdpr assessment. `https://www.gdprbenchmark.com`. Last accessed on 2017-12-10.

[35] Franklin Miller and Alan Wertheimer. *The Ethics of Consent*. Oxford University Press, oct 2009.

[36] Article 29 Data Protection Working Party. `http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/index_en.htm`. Last accessed on 2018-12-22.

[37] Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent. `http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf`, 2011. Last accessed on 2017-11-08.

[38] Article 29 Data Protection Working Party. Guidelines on consent under regulation 2016/679. `http://ec.europa.eu/newsroom/just/document.cfm?doc_id=48849`, 2017. Last accessed on 2017-11-08.

[39] Leigh Ronczka. 8 ways that aligning with iso 27001 help you comply with the gdpr. https://www.pivotpointsecurity.com/blog/how-iso-27001-supports-gdpr-compliance/. Last accessed on 2018-06-06.